

WHAT IS CLAIMED IS:

1                   1.     A method for converting text to concatenated voice by  
2     utilizing a digital voice library and a set of playback rules, the digital voice library  
3     including a plurality of speech items including words and syllables and a  
4     corresponding plurality of voice recordings wherein each speech item corresponds  
5     to at least one available voice recording, the method comprising:

6                   training the digital voice library to associate each syllable speech item  
7     with a literal text syllable of the particular syllable speech item.

1                   2.     The method of claim 1 further comprising:  
2                   receiving a sequence of words including known words that  
3     correspond to word speech items in the digital voice library and including unknown  
4     words;

5                   converting each known word into a word speech item in accordance  
6     with the digital voice library; and

7                   for each unknown word, parsing the unknown word to determine a  
8     sequence of literal text syllables and converting the text syllable sequence to a  
9     sequence of syllable speech items in accordance with the digital voice library.

1                   3.     The method of claim 2 further comprising:  
2                   converting the sequence of word speech items and syllable speech  
3     items into a sequence of voice recordings in accordance with the set of playback  
4     rules.

1                   4.     The method of claim 3 further comprising:  
2                   generating voice data based on the sequence of voice recordings by  
3     concatenating adjacent recordings in the sequence of voice recordings.

1                   5.     The method of claim 4 wherein training the digital voice  
2     library further comprises:  
3                   utilizing a neural network having an input and an output to train the  
4     digital voice library with the neural network receiving the literal text syllable of the

5 particular syllable speech item as input and with the neural network outputting the  
6 associated syllable speech item.

1                   6.     The method of claim 4 wherein training the digital voice  
2 library further comprises:  
3                   manually associating each syllable speech item with the literal text  
4 syllable of the particular syllable speech item.

1                   7.     The method of claim 4 wherein, for each unknown word,  
2 parsing and converting further comprises:  
3                   parsing the unknown word to determine a sequence of literal text  
4 syllables and known words, and converting the sequence to a sequence of syllable  
5 speech items and word speech items in accordance with the digital voice library.

1                   8.     The method of claim 7 wherein parsing further comprises:  
2 parsing the unknown word in the forward direction to determine any  
3 known words;  
4 parsing the unknown word in the reverse direction to determine any  
5 known words;  
6 where any known words overlap, selecting the larger word;  
7 parsing the unknown word in the forward direction to determine any  
8 literal text syllables; and  
9 parsing the unknown word in the reverse direction to determine any  
10 literal text syllables.

1                   9.     The method of claim 7 wherein multiple voice recordings that  
2 correspond to a single speech item represent various inflections of that single speech  
3 item, and wherein converting the sequence of word speech items and syllable speech  
4 items further comprises:  
5 determining a desired inflection for each speech item in the sequence  
6 of speech items based on the set of playback rules; and  
7 determining a sequence of voice recordings by determining a voice  
8 recording for each speech item based on the desired inflection for the particular

9 speech item and based on the available voice recordings that correspond to the  
10 particular speech item.

1 10. The method of claim 7 wherein multiple voice recordings that  
2 correspond to a single speech item represent various inflections and ligatures of that  
3 single speech item, and wherein converting the sequence of word speech items and  
4 syllable speech items further comprises:

5 determining a desired inflection and desired ligatures for each speech  
6 item in the sequence of speech items based on the set of playback rules; and

7 determining a sequence of voice recordings by determining a voice  
8 recording for each speech item based on the desired inflection and desired ligatures  
9 for the particular speech item and based on the available voice recordings that  
10 correspond to the particular speech item.

1 11. The method of claim 4 comprising:

2 for each unknown word, after the unknown word is parsed, storing  
3 results of the parsing in the digital voice library so that a next encounter with the  
4 same unknown word may be handled more efficiently.